Image Sensor Based Visible Light Communication and Its Application to Pose, Position, and Range Estimations

Takaya YAMAZATO[†a] *and* **Shinichiro HARUYAMA**[††], *Fellows*

SUMMARY   This study introduces an image sensor based visible light communication (VLC) and its application to pose, position, and range estimations. There are two types of visible-light receiver: a photodiode and an image sensor. A photodiode is usually used as a reception device of VLC, and an image sensor consisting of a large number of pixels can also be used as a VLC reception device. A photodiode detects the signal intensity of incoming light, while an image sensor not only detects the incoming signal intensity but also an accurate angle of arrival of light emitted from a visible light transmitter such as a white LED light. After angles of arrival of light are detected by an image sensor, positioning and data reception can be performed. The ability of an image sensor to detect an accurate angle of arrival will provide attractive applications of VLC such as pose, position calculation, and range estimation. Furthermore, because the image sensor has the ability to spatially separate sources, outdoor positioning even with strong sunlight is possible by discarding the associated pixels of noise sources.
key words: visible light communication, image sensor, pose estimation, positioning, range estimation

1. Introduction

Visible light communication (VLC) has been recently developed as a method of optical wireless communication that uses solid-state lighting devices such as semiconductor light-emitting diodes (LEDs). Because LEDs can be modulated at high speeds that are undetectable to the human eye, VLC provides the dual functionality of lighting and communication [1], [2].

Growing research trend of VLC may be verified from the numbers of published paper related to VLC. In IEEE Xplore, 53 (2013), 29 (2012) and 13 (2011) journal papers are found and 3 (2013), 5 (2012), and 1 (2011) in IEICE I-Scover. Optical Wireless workshop has been successfully held since 2011 in conjunction with IEEE GLOBECOM, and the special issues on VLC were published in IEEE Communication Magazine in Dec. 2013 and again in July 2014.

VLC standards are already available. The first VLC specification standard was released from JEITA in conjunction with Visible Light Communication Consortium (VLCC) in 2007. Later in 2008, VLCC released their first specification standard adopting and expanding the IrDA physical layer. IrDA transceiver transmission wavelength (IrDA physical layer) is expanded to accommodate the visible light wavelength. By adapting this specification, the existing IrDA optical modules can be used for VLC data transmission with slight modification. In January 2009, IEEE 802.15 (Wireless Personal Area Networks) Task Group 7 (Visible Light Communication) has been launched and in 2011, they have completed a PHY and MAC standard for VLC [3]. Later in July 2013, a study group for camera communications was formed in IEEE 802.15, aiming to standardize VLC for uni-directional, low data rate, short message broadcasts via LEDs using an image sensor as the receiving device.

VLC is still in the introductory phase, but some unique products are already released, such include Picapicamera by CASIO [4] and Underwater VLC Device by MCR [5]. Commercializing VLC technology are undertaken by many research centers, institutes and consortiums, in Europe, US and in Japan, these include VLCC, LiFi Consortium, Smart Lighting Engineering Research Centre, UC-Light Center [6].

The reason for such trend is that VLC provides several advantages [7]. Because the transmitters, or LED light sources, are designed for lighting purposes, they usually have high radiation power. In other words, the signal-to-noise ratio is high for VLC. The visible light spectrum extends from 400 to 790 THz, and it is not regulated globally. The large bandwidth (390 THz) provides opportunities for VLC applications.

It is usually accepted that VLC attains data rates of several Mb/s using a simple on-off keying modulation and a single-element photodiode as a reception device. Such may be enough for VLC startup applications. The data rate may exceed 1 Gb/s by applying more spectrally efficient modulation formats [8]. Such high speed VLC suggests that VLC is a strong candidate to resolve a "spectrum deficit" stated by the Federal Communications Commission, especially for indoor wireless usage [9].

Positioning or localization, also known as visible light positioning, is another attractive application of VLC [10]. Though global navigation satellite systems such as GPS, GLONASS, Galileo, and Compass are widely used as navigation systems for cars and smartphones, their signals from satellites are unable to reach indoors, resulting in an inability to detect the positions of users or objects indoors. Conversely, LEDs that are placed as a ceiling light provide a small personal optical cell, and the isolated LED link delivers a simple positioning solution by sending positioning information from the LED.

In this method to detect position, received signal strength (RSS) of a photodiode is used as a sensor input. Its positioning accuracy by RSS is around several meters. Other methods to detect position using a photodiode include time of arrival, time difference of arrival, phase of arrival, and phase difference of arrival. However, for phosphor-coated LEDs, localization accuracy is limited by the signal bandwidth; the modulation bandwidth of the broad-spectrum component is approximately 2 MHz, while for the blue component, it may be equal to 20 MHz.

The above limitations are due to the use of a photodiode as a receiver, because a photodiode receives incoming light and detects its intensity; however it cannot detect the angle of arrival (AOA) of the incoming light.

Alternatively, an image sensor is able to detect the AOA in addition to the intensity of the incoming light. By using image processing to detect and recognize objects, positioning can be integrated. For example, positioning and data reception are easily achieved. It is reported that VLC can achieve centimeter-level accuracy of indoor positioning at low cost [11]. Furthermore, as the image sensor has the ability to spatially separate sources, outdoor positioning is possible by discarding associated pixels of noise sources.

Although image senor has been used in Infrared communications, the image size is small and the sensor is mainly designed for communication use [12]. It is not suitable for positioning application. We note, however, that the image sensor used for infrared photography can be used.

In this study, we introduce image sensor based VLC and its application to pose, position, and range estimation. After starting with an overview of image sensor based VLC in Sect. 2, we introduce the basics of sensor pose estimation and position measurement in Sect. 3. In Sect. 4, we introduce three unique applications: pose estimation using lights, application to bridge-shape monitoring, and range estimation using phase-only correlation (POC). Finally, in Sect. 5, concluding remarks are presented.

2. Image Sensor Based VLC

An image sensor is a device that converts an optical image into an electronic signal. Image sensors are used in digital cameras, camera modules, video recorders, and other imaging devices. An image sensor consists of $n \times m$ pixels ranging from 320×240 (QVGA) to $157{,}000 \times 18{,}000$ (line scanner). Each pixel contains a photodetector or photodiode and devices for readout circuits. The readout circuits included in pixel devices determine the sensor conversion gain, which is the output voltage per electron collected by photodiode. The readout speed determines the frame rate, typically 30 frames per second. High frame rates are required for many industrial and measurement applications.

A photodiode is usually used as a reception device of VLC, and an image sensor consisting of various pixels can also be used as a VLC reception device [1], [2]. A specific advantage of the image sensor usage is its ability to spatially separate sources, because of the massive number of pixels
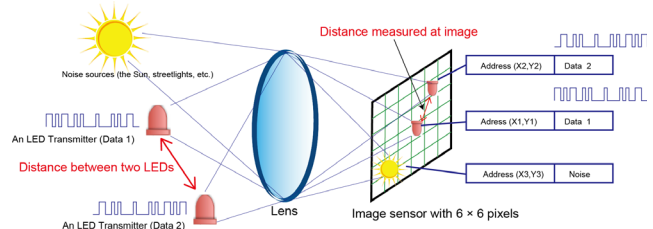

Fig. 1 Image sensor based VLC and positioning.

available.

The ability to spatially separate sources also provides an additional feature to VLC, specifically, the ability to receive and process multiple transmitting sources. Figure 1 shows that the data transmitted from two different LED transmitters can be captured simultaneously.

The output of the image sensor forms a digital electronic representation of the scene. Figure 1 shows that the relative position of two LEDs is projected onto the image sensor. Thus, if the real positions of two LEDs as well as the distance between them are available at the receiver, positioning of the receiver can be performed by triangulation. While such positioning can be processed by an image sensor only (without an aid of VLC), we point out the following advantages of positioning by VLC.

1) Because LEDs can transmit identification numbers (IDs) and other necessary information, no preinstalled data are necessary.
2) Object recognition is much easier because the fast-blinking feature of LEDs is rare in the natural world; thus, no complex image processing is necessary. Fast object recognition is possible even in real-time and receiver-moving (vehicle-driving) cases.
3) The ubiquity of LEDs yields much available data to perform localization. Thus, by using only available relevant data and discarding extraneous data, accuracy can be improved easily.

3. Basics of Sensor Pose Estimation and Position Measurement

In image sensor based VLC, it is possible to compute a sensor (camera) pose with computer vision techniques. This is one of the main advantages compared with other communication techniques. Here, we introduce the basis of computer vision for pose estimation and a pose-estimation method combined with computer vision as well as VLC techniques.

3.1 Single-View Geometry

Computer vision is the study of analyzing the geometry of the real world from images captured with a camera. In general, research issues can be classified into the estimation of the pose of a camera and the reconstruction of the 3D shape of an object (also called 3D modeling) from an image or an
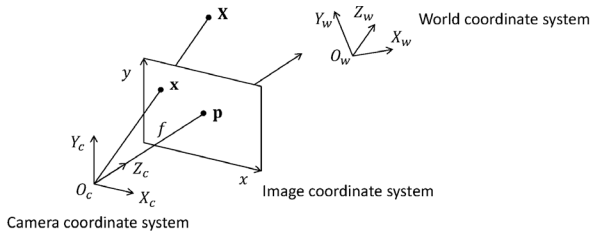
Fig. 2 Single-view geometry.

image sequence. The related research fields of computer vision include image processing, pattern recognition, and machine learning for scene recognition and understanding.

In image sensor based VLC, an image sensor can receive data simultaneously transmitted from multiple lights because each pixel on the sensor is treated as one receiver. Also, the positions of the lights on the image sensor are acquired as feature points and can be used for camera pose estimation. Extensive literature already exists on camera pose estimation using points [13] because it is a traditional research issue. Here, we introduce the basis of single-camera geometry for camera pose estimation. Note that the term "pose" represents the position and orientation relative to some coordinate system in the field of computer vision.

As shown in Fig. 2, camera geometry is described with three coordinate systems: 3D world coordinate system, 2D image coordinate system, and 3D camera coordinate system. A camera pose is usually defined as the position and orientation of the camera coordinate system relative to the world coordinate system. The mathematical representation of the position and orientation is equivalent to the parameters of geometric transformation from the world coordinate system to the camera coordinate system as

$$\tilde{\mathbf{X}}_{\mathbf{c}} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \tilde{\mathbf{X}}_{\mathbf{w}} \tag{1}$$

where $\tilde{\mathbf{X}}_{\mathbf{c}} = (X_c, Y_c, Z_c, 1)^{\mathbf{T}}$ is a homogeneous camera coordinate, $\tilde{\mathbf{X}}_{\mathbf{w}} = (X_w, Y_w, Z_w, 1)^{\mathbf{T}}$ is a homogeneous world coordinate, $\mathbf{R}$ is a 3×3 rotation matrix (orientation), and $\mathbf{t}$ is a 3×1 translation vector (position). Therefore, a camera pose is equivalent to $[\mathbf{R}|\mathbf{t}]$.

The camera coordinate system is defined as the system in which its origin is located at the camera center and the direction of Z_c is perpendicular to the image plane from the camera center. The point of intersection of the image plane and the Z_c axis is called the principal point $\mathbf{p} = (p_x, p_y)$. In the pinhole camera model, a 3D point $\mathbf{X}_{\mathbf{c}} = (X_c, Y_c, Z_c)^{\mathbf{T}}$ in the camera coordinate system is projected onto a 2D point $\mathbf{x} = (x, y)^{\mathbf{T}}$ in the image coordinate system as

$$(x, y)^{\mathbf{T}} = \left(f\frac{X_c}{Z_C} + p_x, f\frac{Y_C}{Z_c} + p_y \right)^{\mathbf{T}} \tag{2}$$

where f is the focal length of a lens. By making a camera calibration matrix $\mathbf{A}$ as
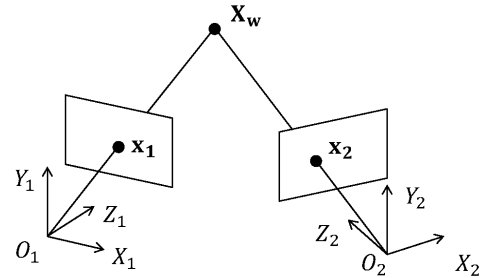
Fig. 3 Two-view geometry.

$$\mathbf{A} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix}. \tag{3}$$

The projection of a 3D point in the world coordinate system onto a 2D point in the image coordinate system is described as

$$\tilde{\mathbf{x}} \sim \mathbf{A} [\mathbf{R}|\mathbf{t}] \tilde{\mathbf{X}}_{\mathbf{w}} \tag{4}$$

where $\tilde{\mathbf{x}} = (x, y, 1)$ is a homogenous image coordinate. This equation is also simplified as

$$\tilde{\mathbf{x}} \sim \mathbf{P}\tilde{\mathbf{X}}_{\mathbf{w}} \tag{5}$$

$$\mathbf{P} = \mathbf{A} [\mathbf{R}|\mathbf{t}] \tag{6}$$

where $\mathbf{P}$ is a 3×4 perspective projection matrix that also represents a camera pose.

To estimate a camera pose by solving the above equations, attaining multiple sets of $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{X}}_{\mathbf{w}}$ is mandatory. For example, $\mathbf{P}$ is linearly computed from six sets because there are 12 unknown parameters in $\mathbf{P}$ and two equations are prepared from one set. If $\mathbf{A}$ is known using a camera calibration technique [14], it is possible to compute a camera pose up to four pairs of solutions [15]. Because many solutions under different conditions have been proposed in the literature, the solutions are not limited to the above ones.

3.2 Basis of Triangulation

Triangulation is a method for 3D position measurement of an object in the real world with two known cameras. A 3D point in the world coordinate system is projected onto a 2D point in the image coordinate system. When a 3D point $\tilde{\mathbf{X}}_{\mathbf{w}}$ is projected onto two cameras, this is mathematically described as

$$\tilde{\mathbf{x}}_{\mathbf{1}} \sim \mathbf{P}_{\mathbf{1}}\tilde{\mathbf{X}}_{\mathbf{w}} \tag{7}$$

$$\tilde{\mathbf{x}}_{\mathbf{2}} \sim \mathbf{P}_{\mathbf{2}}\tilde{\mathbf{X}}_{\mathbf{w}} \tag{8}$$

where $\mathbf{P}_{\mathbf{i}}$ is a perspective projection matrix of each camera and $\tilde{\mathbf{x}}_{\mathbf{i}}$ is a homogeneous image coordinate computed by projecting $\tilde{\mathbf{X}}_{\mathbf{w}}$ onto each image plane with $\mathbf{P}_{\mathbf{i}}$. $\tilde{\mathbf{x}}_{\mathbf{1}}$ and $\tilde{\mathbf{x}}_{\mathbf{2}}$ are considered as correspondences between the two images. The use of two known cameras means that $\mathbf{P}_{\mathbf{1}}$ and $\mathbf{P}_{\mathbf{2}}$ are known. This is also graphically explained in Fig. 3. In triangulation,
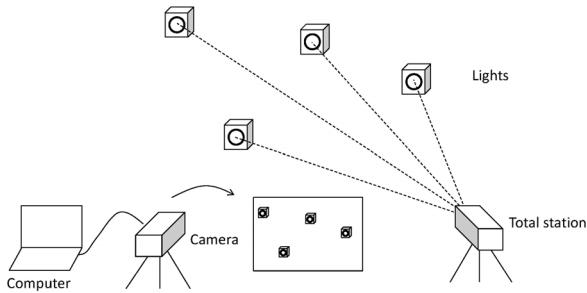
Fig. 4   Pose estimation using lights.

Table 1   List of IDs and world coordinates.

ID	world coordinate
1	(X_1, Y_1, Z_1)
2	(X_1, Y_1, Z_1)
⋮	⋮
N	(X_N, Y_N, Z_N)

Fig. 5   Bridge shape measurement.

Fig. 6   Snapshot of a bridge being measured.

the goal is to compute $\tilde{\mathbf{X}}_\mathbf{w}$ by finding $\tilde{\mathbf{x}}_1$ and $\tilde{\mathbf{x}}_2$.

4.   Application to Pose, Position, and Range Estimations

4.1   Pose Estimation Using Lights

It is necessary to acquire multiple sets of a world coordinate and its projected image coordinate for camera pose estimation. In image sensor based VLC, such sets are acquired using lights.

Figure 4 shows an overview of pose estimation using lights [16]. Lights are first placed in a target scene and their 3D world coordinates are measured using an electronic distance meter such as a total station. To compute the image coordinate of a light and receive its transmitted data, a camera captures the lights at a fixed position because a light should be captured at the same position through an image sequence.

When the transmitted data can include the world coordinate of the light, the set of the world coordinate of a light and its projected image coordinate is simply acquired. However, the size of transmitted data may not be adequate. Therefore, the transmitted data can be just an ID and the list of the IDs of lights and their world coordinates, as shown in Table 1, should be stored in a computer before. If the number of acquired sets satisfies the condition of pose estimation, a camera pose can be computed. Note that the accuracy of an estimated camera pose depends on the accuracy of the measurement of 3D world coordinates. In the following subsection, we explain an efficient method for computing the image coordinate of a blinking light.

4.2   Application to Bridge Shape Monitoring

Image sensor based VLC works under different types of illumination, from daytime to midnight, because a light actively
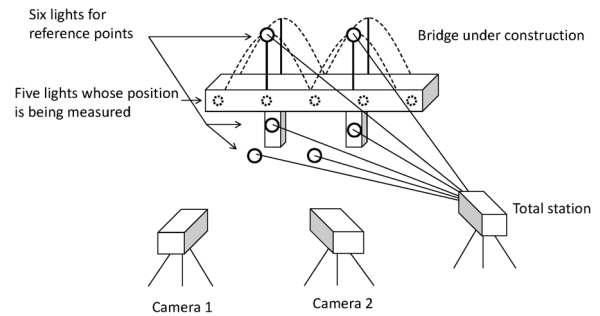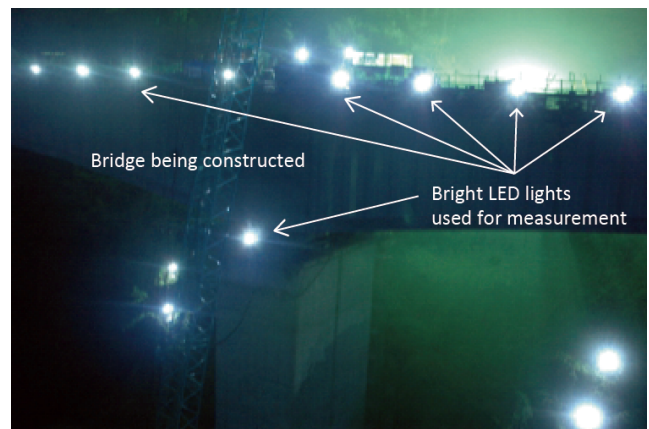
transmits data by blinking. This feature is convenient for bridge shape monitoring. In general, the shape of a bridge is deformed because the fabric of the bridge dilates and shrinks according to the ambient temperature. When a bridge is built, such deformation should be monitored for the early detection of problems.

Figure 5 shows an overview of bridge shape monitoring. On the bridge, two different types of lights are placed: points for measuring the deformation and points for camera pose estimation. The latter points are typically called reference points and should be placed at the points where deformation does not occur.

The monitoring procedure is as follows. From the reference points, the poses of Camera 1 and Camera 2 are first computed. At each camera, the image coordinates of lights and the transmitted data are then computed. By selecting the pixels that send the same data, the corresponding pixels of a light in two images are computed. Such pixels are triangulated to compute the 3D position of the light. Figure 6 shows a snapshot of a bridge under construction whose positions are being measured using the technology described here. The positions of the bright LED lights on the bridge are being measured, while the positions of the bright LED lights on the ground and at the pier are measured by a total station before.

Fig. 7 1024 LED array. Each LED is arrange in square matrix of 32×32. The size of the LED array is 465 mm × 465 mm, with the LED spacing of 15 mm. The LEDs are the same as the actual LED traffic lights in Japan.

Fig. 8 Captured LED overlap with multiple pixels.

Fig. 9 POC image G_{12} for $f_1(m, n)$ and $f_2(m, n) = f_1(m + \tau_m, n + \tau_n)$.

4.3 Range Estimation Using Phase-Only Correlation

Normally, when we use triangulation, either two transmitters and one camera, or one transmitter and two cameras, are used. The longer the distance between two transmitters or two cameras, the better range estimation is obtained. Here, we consider range estimation using two LEDs placed on the edges of an LED array and one camera receiver.

Figure 7 shows the LED array used for the experiment. Although estimating from an LED array is attractive in some applications, the breadth of LED array is small (465 mm). Therefore, poor range accuracy is expected because the distance measured at the camera coordinate only considers a small value. In fact, for the LED array shown in Fig. 7, the projected width of the LED array is only 9 pixels at 60 m using a 35 mm lens. In such cases, the range error may reach 3.5 m at a distance of 60 m.

A solution is to estimate the projected width of the LED array in subpixel resolution using POC [17], [18].

Before we discuss the details of POC, we begin with a mathematical equation for estimating range from the projected width of the LED array.

Let L be a range to be estimated, then we obtain L as [19]

$$L = \frac{fS}{W\alpha}, \tag{9}$$

where S is the width of the LED array, α is a pixel size of the image sensor in μm, and f is the focal distance of the camera lens. The projected width of the LED array, denoted by W, is the parameter to estimate. Because S, α, and f are constants, the accuracy of the distance is controlled by W.

Figure 8 illustrates 3×3 pixels around an LED. The LED overlaps with six pixels. We estimate the position of the LED as the highest luminance pixel in the six pixels. The estimated center of the LED is considered as the center of pixel shown by a circle, whereas the actual center of the LED is shown by a triangle. As shown, the actual center of
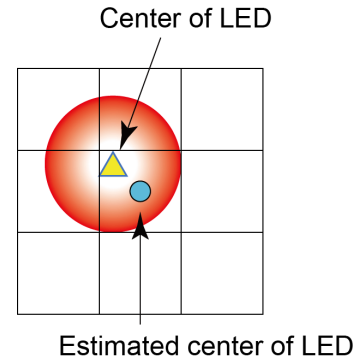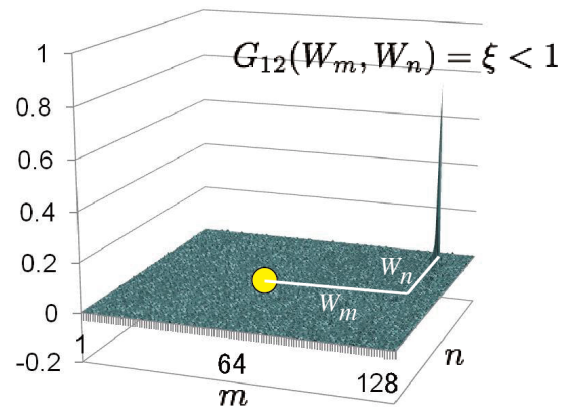
the LED is different from the estimated center.

4.3.1 Subpixel Range Estimation Using POC

POC is a pattern-matching algorithm and defined as a modified correlation algorithm such that the amplitude components of the Fourier-transformed images are replaced with a constant. In general, POC is used for position displacement detection of the same object.

In POC, the position displacement of the two images is reflected by cross-correlation. Let us consider $f_2(m, n) = f_1(m+\tau_m, n+\tau_n)$ that the original image $f_1(m, n)$ is shifted to τ_m in the m direction and τ_n in the n direction, respectively, where τ_m and τ_n are consecutive values. Note that pixel-width W takes a natural number. If we denote δ_m and δ_n as a small discrete value, τ_m and τ_n are given by

$$\tau_m \cong W_m + \delta_m, \tag{10}$$
$$\tau_n \cong W_n + \delta_n \tag{11}$$

Then we obtain POC as the following [19]:

$$G_{12}(r - \tau_m, s - \tau_n)$$
$$\cong \begin{cases} 1 & r = W_m + \delta_m, s = W_n + \delta_n \\ \xi(< 1) & r = W_m, s = W_n \\ \varepsilon & r \neq W_m, s \neq W_n \end{cases} \tag{12}$$

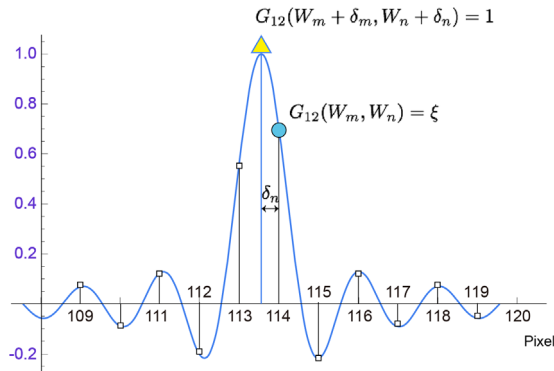Figure 9 shows a result of the POC using two images.

Fig. 10 Sinc function approximation of G12(r, s) shown in cross-section of Fig. 8 in n direction.

Table 2 Parameters of high-speed camera (Photron FASTCAM-1024PCI 100 K).

Pixel size α	17 μm
Focal distance f	35 mm
Lens diaphragm	8
Focus of a lens	Infinity
Resolution $(P \times P)$	128 × 128 pixel

Unfortunately, POC output is a function of the integer values of W_m and W_n, respectively, therefore the accuracy is still limited to the pixel width. Because the actual values of displacement are denoted by $\tau_m \cong W_m + \delta_m$ and $\tau_n \cong W_n + \delta_n$ it is essential to estimate δ_m and δ_n.

From (11), if we know δ_m and δ_n, then $G_{12}(r, s)$ becomes one. Moreover, we can approximate $G_{12}(r, s)$ by the sinc function.

$$G_{12}(r, s) \cong \operatorname{sinc}(r + \delta_m)\operatorname{sinc}(s + \delta_n) \qquad (13)$$

The peak of $G_{12}(r, s)$ can be obtained by varying δ_m and δ_n. The center of the LED is given by

$$(W_m + \hat{\delta}_m, W_n + \hat{\delta}_n) = \underset{-1 \le \delta_m \le 1, -1 \le \delta_n \le 1}{\arg\max} G_{12}(r, s), \qquad (14)$$

Where $\hat{\delta}_m$ and $\hat{\delta}_n$ are the estimates of δ_m and δ_n, respectively. Figure 10 shows a cross-section image of Fig. 9 in the n direction. The triangle in Fig. 8 expresses the center of the LED.

4.3.2 Experimental Results

We performed experimental measurements in 5 m increments in a static condition, for distances from 20 m to 60 m. The LED array is shown in Fig. 7, and Table 2 summarizes the parameters of the high-speed camera.

Figure 11 illustrates the average range estimation error levels. The horizontal axis represents the distance between the LED array and the receiver camera, and the vertical axis represents the range estimation errors. The graph label "pixel level estimation" represents the conventional range estimation method, whose accuracy is limited to pixel
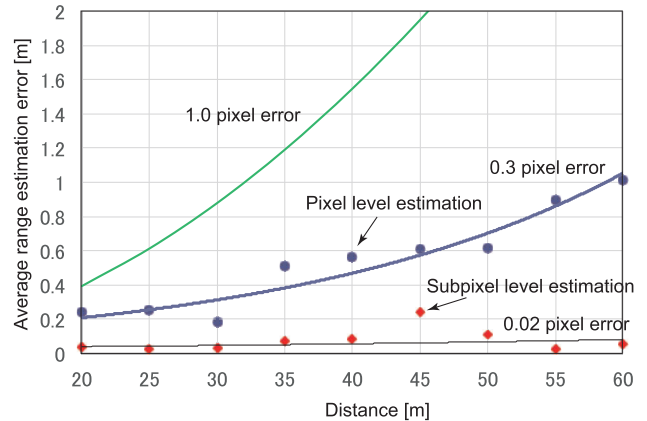

Fig. 11 Average range estimation error versus distance.

width. The graph label "subpixel level estimation" represents the range estimation error levels using POC, whose accuracy is at the subpixel level. As we confirm from the figure, the curve of "subpixel level estimation" achieves less than 0.3 m estimation error from 20 to 60 m. We note that this corresponds to a 0.02-pixel subpixel estimation accuracy. In contrast, the curve of the "pixel level estimation" scheme shows more than 1 m estimation error at 60 m.

5. Conclusions

In this study, we introduced the image sensor based VLC and its application to pose, position, and range estimations. From the basics of sensor pose estimation and position measurement, we showed application examples of accurate sensor pose estimation, position measurement for civil engineering, and range estimation using POC.

An image sensor is able to spatially separate visible light sources, receive and demodulate optical signals from a visible light transmitter at a pixel position where the optical light is projected, and detect accurate arrival angle of an incoming light. By using the unique characteristics of image sensors, applications shown in this study are made possible. For example, the capability of accurate arrival angle detection makes it possible to perform accurate sensor pose estimation and position measurement for civil engineering, and the capability of image processing makes it possible to perform range estimation.

References

[1] S. Haruyama and T. Yamazato, "[Tutorial] Visible light communications," IEEE International Conference on Communications, June 2011.

[2] T. Yamazato, I. Takai, H. Okada, T. Fujii, T. Yendo, S. Arai, M. Andoh, T. Harada, K. Yasutomi, K. Kagawa, and S. Kawahito, "Image sensor based visible light communication for automotive applications," IEEE Commun. Mag., vol.52, no.7, pp.88–97, July 2014.

[3] S. Rajagopal, R.D. Roberts, and S.-K. Lim, "IEEE 802.15.7 visible light communication: Modulation schemes and dimming support," IEEE Commun. Mag., pp.72–82, March 2012.

[4] http://www.casio-isc.com/en/

[5] http://www.mcrvlc.jp/?lang=english

[6] http://www.marketsandmarkets.com/Market-Reports/visible-light-communication-market-946.html

[7] Harald Haas, "Wirelss data from every light bulb," http://www.ted.com/talks/harald_haas_wireless_data_from_every_light_bulb

[8] M. Ishida, S. Miyauchi, T. Komine, S. Haruyama, and M. Nakagawa, "An architecture for high-speed parallel wireless visible light communications system using 2D image sensor and LED transmitter," Proc. International Symposium on Wireless Personal Multimedia Communications (WPMC'05) (CD-ROM), pp.1523–1527, 2005.

[9] L. Grobe, A. Paraskevopoulos, J. Hilt, D. Schulz, F. Lassak, F. Hartlieb, C. Kottke, V. Jungnickel, and K. Langer, "High-speed visible light communication systems," IEEE Commun. Mag., pp.60–66, Dec. 2013.

[10] J. Armstrong, Y. Sekercioglu, and A. Neild, "Visible light positioning: A roadmap for international standardization," IEEE Commun. Mag., pp.68–73, Dec. 2013.

[11] A. Jovicic, J. Li, and T. Richardson, "Visible light communication: Opportunities, challenges and the path to market," IEEE Commun. Mag., pp.26–32, Dec. 2013.

[12] J.M. Kahn, R. You, A.G. Weisbin, B.K. Teik, and A. Tang, "Imaging diversity receivers for high-speed infrared wireless communication," IEEE Commun. Mag., pp.88–94, Dec. 1998.

[13] R. Szeliski, Computer vision: Algorithm and applications, Springer, 2011.

[14] Z. Zhang, "A flexible new technique for camera calibration," IEEE Trans. Pattern. Anal. Mach. Intell., pp.1330–1334, 2000.

[15] D. Nister, "A minimal solution to the generalised 3-point pose problem," Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.560–567, 2004.

[16] H. Uchiyama, M. Yoshino, H. Saito, M. Nakagawa, S. Haruyama, T. Kakehashi, and N. Nagatomo, "Photogrammetric system using visible light communication," Proc. 34th Annual Conference of the IEEE Industrial Electronics Society, pp.1771–1776, 2008.

[17] K. Takita, T. Aoki, Y. Sasaki, T. Higuchi, and K. Kobayashi, "High-accuracy subpixel image registration based on phase-only correlation," IEICE Trans. Fundamentals, vol.E86-A, no.8, pp.1925–1933, Aug. 2003.

[18] H. Foroosh, J.B. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," IEEE Trans. Image Process., vol.11, no.3, pp.188–200, March 2002.

[19] A. Ohmura, T. Yamazato, H. Okada, T. Fujii, T. Yendo, and S. Arai, "Accuracy improvement by phase only correlation for distance estimation scheme for visible light communications using an LED array and a high-speed camera," Proc. 20th World Congress on Intelligent Transport Systems, Oct. 2013.

Takaya Yamazato is a professor at the Institute of Liberal Arts and Sciences, Nagoya University, Japan. He received the Ph.D. degree from Department of Electrical Engineering, Keio University, Yokohama, Japan, in 1993. From 1993 to 1998, he was an Assistant Professor in the Department of Information Electronics, Nagoya University, Japan. From 1997 to 1998, he was a visiting researcher of the Research Group for RF Communications, Department of Electrical Engineering and Information Technology, University of Kaiserslautern. In 1998, he gave a 1/2-day tutorial entitled "Introduction to CDMA ALOHA" at Globecom held in Sydney Australia. Since then, he has been serving as a TPC member of Globecom and ICC. In 2006, he received the IEEE Communication Society 2006 The Best Tutorial Paper Award. He served as a co-chair of Wireless Communication Symposia of ICC2009 and he was a co-chair of Selected Areas in Communication Symposia of ICC2011. From 2009–2010, he was an editor in chief of Japanese Section of IEICE Transaction on Communications. His research interests include visible light communication, satellite and mobile communication systems, and ITS.

Shinichiro Haruyama is a professor at The Graduate School of System Design and Management, Keio University, Yokohama, Japan. He received an M.S. in engineering science from University of California at Berkeley in 1983 and a Ph.D. in computer science from the University of Texas at Austin in 1990. He worked for Bell Laboratories of AT&T and Lucent Technologies, U.S.A. from 1991 to 1996, and for Sony Computer Science Laboratories, Inc. from 1998 to 2002. He then joined Department of Information and Computer Science, Faculty of Science and Technology, Keio University in 2002, and moved to a The Graduate School of System Design and Management, Keio University in 2008. He is a chair of Visible Light Communication Association. He is a fellow of IEICE. His research interests include visible light communication, wireless communication, system design automation, and integrated circuit design.